

# Classroom assessment and the *National Science Education Standards*

Committee on Classroom  
Assessment and the  
National Science  
Education Standards

J. Myron Atkin,  
Paul Black,  
Janet Coffey, Editors

Center for Education

Division of Behavioral  
and Social Sciences  
and Education

National Research Council

## 3

### Assessment in the Classroom

The primary audiences for this chapter are classroom teachers and teacher educators. The chapter offers a guiding framework to use when considering everyday assessments and then discusses the roles and responsibilities of teachers and students in improving assessment. Administrators also may be interested in the material presented in this chapter.

Assessment usually conjures up images of an end-of-unit test, a quarterly report card, a state-level examination on basic skills, or the letter grade for a final laboratory report. However, these familiar aspects of assessment do not capture the full extent or subtlety of how assessment operates every day in the classroom. The type of classroom assessment discussed in this chapter focuses upon the daily opportunities and interactions afforded to teachers and students for collecting information about student work and understandings, then uses that information to improve both teaching and learning. It is a natural part of classroom life that is a world away from formal examinations--both in spirit and in purpose.

During the school day, opportunities often arise for producing useful assessment information for teachers and students. In a class discussion, for example, remarks by some of the students may lead the teacher to believe that they do not understand the concept of energy conservation. The teacher decides that the class will revisit an earlier completed laboratory activity and, in the process, examine the connections between that activity and the discussion at hand. As groups of students conduct experiments, the teacher circulates around the room and questions individuals about the conclusions drawn from their data. The students have an opportunity to reflect on and demonstrate their thinking. By trying to identify their sources of evidence, the teacher better understands where their difficulties arise and can alter their teaching accordingly and lead the students toward better understanding of the concept.

As another example, a planning session about future science projects in which the students work in small groups on different topic issues leads to a discussion about the criteria for judging the work quality. This type of assessment discussion, which occurs before an activity even starts, has a

powerful influence on how the students conduct themselves throughout the activity and what they learn. During a kindergarten class discussion to plan a terrarium, the teacher recognizes that one of the students confuses rocks for living organisms and yet another seems unclear about the basic needs of plants. So the conversation is turned toward these topics to clarify these points. In this case, classroom teaching is reshaped immediately as a result of assessments made of the students' understanding.

Abundant assessment opportunities exist in each of these examples. Indeed, Hein and Price (1994) assert that anything a student does can be used for assessment purposes. This means there is no shortage of opportunities, assessment can occur at any time. One responsibility of the teacher is to use meaningful learning experiences as meaningful assessment experiences.

Another is to select those occasions particularly rich in potential to teach something of importance about standards for high-quality work. To be effective as assessment that improves teaching and learning, the information generated from the activity must be used to inform the teacher and/or students in helping to decide what to do next. In such a view, assessment becomes virtually a continuous classroom focus, quite indistinguishable from teaching and curriculum.

The *Standards* convey a view of assessment and learning as two sides of the same coin and essential for all students to achieve a high level of understanding in science. To best support their students' learning, teachers are continuously engaged in ongoing assessments of the learning and teaching in their classroom. An emphasis on formative assessment--assessment that informs teaching and learning and occurs throughout an activity or unit--is incorporated into regular practice. Furthermore, teachers cultivate this integrated view of teaching, learning, and continuous assessment among their students. When formative assessment becomes an integral part of classroom practice, student achievement is enhanced (Black & Wiliam, 1998a; Crooks, 1988; Fuchs & Fuchs, 1986). However, as discussed in the previous chapter, research also indicates that this type of assessment often is not recognized as significant by teachers, principals, parents, or the general public, and is seldom articulated or featured as a priority. [Box 3-1](#) provides definitions for "formative" and "summative," which pertain to the two main functions that assessment can take.

The centrality of inquiry in the vision of science education advanced in the *Standards* provides a particularly compelling reason to take a closer look at classroom assessment, and formative assessment, in particular. If students are to *do* science, not solely verbalize major facts and principles, they should engage in activity that extends over several days or weeks. Their work should be less episodic and fractured than lesson-based science teaching. A different kind of assessment is necessary, one that is designed to help students get better at inquiring into the world of science (NRC, 2000). The best way to support inquiry is to obtain information about students while they are actually engaged in science investigations with a view toward helping them develop their understandings of both subject matter and procedure. The information collected by teachers and students while the students are at work can be used to guide their progress. A teacher asks questions that may help spur thinking

about science concepts that are part of the investigation and may help students understand what it takes to do work that comports with high standards. At the end, the information may be collected and reviewed to form a basis for summative evaluations.

## FEATURES OF FORMATIVE ASSESSMENT

To help design and implement an effective and efficient classroom assessment system, we offer the following general template for designing and integrating formative assessment into regular classroom practice.

- *Where are you trying to go?*
- *Where are you now?*
- *How can you get there?*

Having posed these questions as a guide, it is important to note that no one blueprint or single best model exists for using assessment as a tool that, first and foremost, supports and facilitates student learning. Each teacher needs to develop a system that works for him or her. By making explicit desirable features of assessment, these three critical questions provide a framework for achieving powerful classroom assessment. The questions and the obtained responses are tightly interconnected and interdependent and they are not new. Based on experience, many teachers both intuitively and purposefully consider these questions every day. Attention to them is part of good teaching.

Through the vignettes and discussion that follow, we hope to make features of formative assessment more explicit and, in doing so, highlight how intimately they are connected to teaching.

---

### A Look Inside Two Classrooms

The seventh-grade students in Ms. K's science class are working on long-term research projects investigating their local watershed. In addition to class discussions, laboratory activities, and field trips, small groups of students are exploring various areas of particular interest and importance. One group is surveying local industrial, agricultural, and residential areas to locate general and point sources of pollutants. Another group is examining water quality. A third group is focusing on how the local ecosystem influences water quality. During project work-time, Ms. K conducts conferences with groups of students about their projects. In these small groups, the students share the details of their project; from content to process, Ms. K keeps herself informed on the working status of the different groups. Information she gathers from these conferences feeds into her decisions about allotment of work time, possible resource suggestions, and areas where she can identify additional learning opportunities. She also is able to note progress that occurs throughout the project, as well as from the last time she engaged in a similar activity with students. For example, after one of the discussions, she realized that the students in one group were not connecting algal blooms to possible sources of

pollutants. She asked questions that encouraged them to explore possible causes of the burst in algal blooms, and together they devised an experiment that had the potential of providing them with some useful, additional information.

Journals kept by the students become the stimulus for regular reflections on learning and the connections between their topic to the bigger picture of the local watershed. Ms. K collects the journals weekly. The journal reflections inform her about the progress of the groups and the difficulties they are having, and so serve as a springboard for class discussion. From reading student responses and listening to discussion, Ms. K knows that some of her students are making deeper connections, and many are making different connections. Painting the broad landscape for the entire class will give those who are struggling to find a broader context for their work and sustain their inquiries, so she decides to create an opportunity to do so. When she is not in discussions with students, she mills around the areas where her students work, moving from group to group, sometimes asking questions, sometimes just listening and observing before she joins the next group. She carries a clipboard on which she jots down notes, quotes, and questions that she will want to come back to with a particular student or the entire group. Through the journals, her observations, the discussions, and other assessment activities, Ms. K stays connected to the sense her students are making of their work as it unfolds.

At the very beginning of the project, Ms. K and her students started conversations about how their projects would be assessed. As a class, they cycle back through the criteria that were established, deepening understanding by highlighting exemplars from past projects and just talking through what constitutes quality work. They share examples of visual display boards, written reports, and models from other projects. Ms. K wants to make sure that each student understands the standards that they are expected to meet. Students chose many of the criteria by which they wish their peers to evaluate them, and, with Ms. K's help, they developed an evaluation rubric that will be ready on presentation day--now just 2 weeks away. At that time, they will be making public reports to peers, parents, and community members.

SOURCE: Coffey (2001).



The King School was reforming its science curriculum. After considerable research into existing curriculum materials and much discussion, the team decided to build a technology piece into some of the current science studies. The third-grade teacher on the team, Ms. R., said that she would like to work with two or three of her colleagues on the third-grade science curriculum. They selected three topics that they knew they would be teaching the following year: life cycles, sound, and water.

Ms. R. chose to introduce technology as part of the study of sound. That winter, when the end of the sound study neared, Ms. K., was ready with a new culminating activity--making musical instruments. She posed a question to the entire class: Having studied sound for almost 6 weeks, could they design and

make musical instruments that would produce sounds for entertainment? Ms. R had collected a variety of materials, which she now displayed on a table, including boxes, tubes, string, wire, hooks, scrap wood, dowels, plastic, rubber, fabric and more. The students had been working in groups of four during the sound study, and Ms. R asked them to gather into those groups to think about the kinds of instruments they would like to make. Ms. R asked the students to think particularly about what they knew about sound, what kind of sound they would like their instruments to make, and what kind of instrument it would be. How would the sound be produced? What would make the sound? She suggested they might want to look at the materials she had brought in, but they could think about other materials too.

Ms. R sent the students to work in their groups. Collaborative work had been the basis of most of the science inquiry the student had done; for this phase, Ms. R felt that the students should work together to discuss and share ideas, but she suggested that each student might want to have an instrument at the end to play and to take home.

As the students began to talk in their groups, Ms. R added elements to the activity. They would have only the following 2 weeks to make their instruments. Furthermore, any materials they needed beyond what was in the boxes had to be materials that were readily available and inexpensive.

Ms. R. knew that planning was a challenge for these third graders. She moved among groups, listening and adding comments. When she felt that discussions had gone as far as they could go, she asked each group to draw a picture of the instruments the children thought they would like to make, write a short piece on how they thought they would make them, and make a list of the materials that they would need. Ms. R made a list of what was needed, noted which children and which groups might profit from discussing their ideas with one another, and suggested that the children think about their task, collect materials if they could, and come to school in the next week prepared to build their instruments.

Ms. R. invited several sixth graders to join the class during science time the following week, knowing that the third-grade students might need their help in working with the materials. Some designs were simple and easy to implement, for example, one group was making a rubber-band player by stretching different widths and lengths of rubber bands around a plastic gallon milk container with the top cut off. Another group was making drums of various sizes using some thick cardboard tubes and pieces of thin rubber roofing material. For many, the designs could not be translated into reality, and much change and trial and error ensued. One group planned to build a guitar and designed a special shape for the sound box, but after the glued sides of their original box collapsed twice, the group decided to use the wooden box that someone had added to the supply table. In a few cases, the original design was abandoned, and a new design emerged as the instrument took shape.

At the end of the second week, Ms. R set aside 2 days for the students to reflect on what they had done individually and as a class. On Friday, they were once again to draw and write about their instruments. Where groups had worked together on an instrument, one report was to be prepared. On the next



Monday, each group was to make a brief presentation of the instrument, what it could do, how the design came to be, and what challenges had been faced. As a final effort, the class could prepare a concert for other third grades.

In making the musical instruments, students relied on knowledge and understanding developed while studying sound, as well as the principles of design, to make an instrument that produced sound. The assessment task for the musical instruments follows. The titles emphasize some important components of the assessment process.

**Science Content:** The K-4 science content standard on science and technology is supported by the idea that students should be able to communicate the purpose of a design. The K-4 physical science standard is supported by the fundamental understanding of the characteristics of sound, a form of energy.

**Assessment Activity:** Students demonstrate the products of their design work to their peers and reflect on what the project taught them about the nature of sound and the process of design.

**Assessment Type:** This can be public, group, or individual, embedded in teaching.

**Assessment Purpose:** This activity assesses student progress toward understanding the purpose and processes of design. The information will be used to plan the next design activity. The activity also permits the teacher to gather data about understanding of sound.

**Data:** Observations of the student performance.

**Context:** Third-grade students have not completed a design project. Their task is to present the product of their work to their peers and talk about what they learned about sound and design as a result of doing the project. This is a challenging task for third-grade students, and the teacher will have to provide considerable guidance to the groups of students as they plan their presentations.

As described in the science standards, the teacher provided the following directions that served as a framework that students could use to plan their presentations.

1. Play your instrument for the class.
2. Show the class the part of the instrument that makes the sound.
3. Describe to the class the purpose (function) that the other parts of the instrument have.
4. Show the class how you can make the sound louder.
5. Show the class how you can change the pitch (how high or how low the sound is) of the sound.
6. Tell the class about how you made the instrument, including
  - What kind of instrument did you want to make?
  - How like the instrument you wanted to make is the one you actually made?

- Why did you change your design?
  - What tools and materials did you use to make your instrument?
7. Explain why people make musical instruments.

In order to evaluate the student performance, the teacher used the following guidelines:

Student understanding of sound will be revealed by understanding that the sound is produced in the instrument by the part of the instrument that vibrates (moves rapidly back and forth), that the pitch (how high or how low) can be changed by changing how rapidly the vibrating part moves, and the loudness can be changed by the force (how hard you pluck, tap, or blow the vibrating part) with which the vibrating part is set into motion. An average student performance would include the ability to identify the source of the vibration and ways to change either pitch or loudness in two directions (raise and lower the pitch of the instrument or make the instrument louder and softer) or change the pitch and loudness in one direction (make the pitch higher and the sound louder). An exemplary performance by a student would include not only the ability to identify the source of the vibration but also to change pitch and loudness in both directions.

Student understanding of the nature of technology will be revealed by the student's ability to reflect on why people make musical instruments--to improve the quality of life--as well as by their explanations of how they managed to make the instrument despite the constraints faced--that is, the ability to articulate why the conceptualization and design turned out to be different from the instrument actually made. (p. 49)

SOURCE: NRC (1996).

---

There is no one best assessment system for the classroom. What works for Ms. K or Ms. R in their classrooms may not work in another. What is important is that assessment is an ongoing activity, one that relies on multiple strategies and sources for collecting information that bears on the quality of student work and that then can be used to help both the students and the teacher think more pointedly about how the quality might be improved.

In the first vignette, Ms. K is helping her students by painting the broad landscape so that they can see how their work fits into a wider context. She also reminds them of the criteria for quality work. Thus, she is helping them to develop a clear view of what they are to achieve and where they are going. At this stage, the view is usually clearer to the teacher than to the students. One of her responsibilities is to help the students understand and share the goals, which will become progressively clearer to them as the inquiry progresses.

To chart student progress, Ms. K relies on several strategies and sources: observations, conversations, journal assignments, student work, and a final presentation. These opportunities are part of the natural flow of classroom life, indistinguishable for her and for the students from collecting data, discussing findings, planning next steps, drawing conclusions, and communicating findings about the main concepts they are expected to learn. In helping her

students to reach their goal, she bases her actions on multiple pieces of evidence that she gleans from activities embedded in her teaching and curriculum. She uses this information to make decisions about work time, about support she needs to provide, and about resource suggestions.

Ms. R also uses assessment in strategic and productive ways. She frames an assessment task in a way that will engage students to learn as they prepare for the final presentation and concert. Peer-design reviews, conversations, and other assessments were built into the activity of designing and building instruments so that students could draw from these to inform their design and construction of instruments. She provides the students with prompts and elements that should be included in their presentations so that the students will be clear on what is required. She has clear guidelines about the quality and depth of responses in terms of how students will demonstrate their understandings and skills.

The usefulness of assessment does not stop at teachers collecting information in the course of their teaching and providing feedback. Like Ms. K and Ms. R, they plan and structure specific assessment events, such as individual conferences with students, occasions for the students to write about a topic, design reviews, observations of students at work, presentations of work, and initiating whole-class discussion of what they have learned so far. These are just some of the many assessment activities and methods available to teachers and students. In these same scenarios, teachers could also have integrated the use of additional written assessments--including selected response, short answer, essay, lab reports, homework problems, among others--into their teaching in ways that would generate rich assessment opportunities.

Throughout this text, we have attempted to avoid technical terms whenever possible. When we do use them, we try to offer a definition or use it in a context where its meaning makes sense. [Box 3-2](#) provides operational definitions of several terms you will find in the assessment literature.

Now, consider the assessment in the two vignettes in light of the following three guiding questions: Where are you trying to go? Where are you now? How can you get there?

## WHERE ARE YOU TRYING TO GO?

### Clear Criteria

The goals articulated in the *Standards* arise from their emphasis on the active nature of science and their stress on the range of activities that encompass what it means to do science and to understand both specific concepts and the subject area as a whole. Thus, the *Standards* advocate going beyond the coverage of basic facts to include skills and thought processes, such as the ability to ask questions, to construct and test explanations of phenomena, to communicate ideas, to work with data and use evidence to support arguments, to apply knowledge to new situations and new questions, to problem solve and make decisions, and to understand history and nature of scientific knowledge



(NRC, 1996). To best assist students in their science learning, assessment should attend to these many facets of learning, including content understanding, application, processes, and reasoning.

In his book on classroom assessment for teachers, Stiggins (2001) writes,

The quality of any assessment depends first and foremost on the clarity and appropriateness of our definitions of the achievement target to be assessed...We cannot assess academic achievement effectively if we do not know and understand what that valued target is. (p. 19)

As Stiggins states, it is important that teachers have clear performance criteria in mind before they assess student work and responses. Ms. R's guidelines included attention to both: she expected her students to demonstrate an understanding of concepts of sound, such as causes of pitch, as well as the nature of technology. Before the students engaged in the assessment, Ms. R had outlined how she would evaluate the student responses in each area.

Clarity about the overall goals is only a first step. Given that goals are clear, the teacher has to help the students achieve greater clarity. This usually entails identification of somewhat discrete stages that will help the students to understand what is required to move toward the goal. These intermediate steps often emerge as the study progresses, often in lesson design and planning but also on the spot in the classroom as information about the students' levels of understanding become clearer, new special interests become apparent, or unexpected learning difficulties arise. This complex, pedagogical challenge is heightened because the goals that embody the standards and the related criteria need to be understood by *all* students.

One of the goals of the *Standards* is for all students to become independent lifelong learners. The standards emphasize the integral role that regular self-assessment plays in achieving this goal. The document states:

Students need the opportunity to evaluate and reflect on their own scientific understanding and ability. Before students can do this, they need to understand the goals for learning science. The ability to self-assess understanding is an essential tool for self-directed learning. (p. 88)

Sadler (1989) emphasizes the importance of student understanding of what constitutes quality work, "The indispensable condition for improvement is that the student comes to hold a concept of quality roughly equivalent to that held by the teacher..." (p. 121). Yet, conveying to students the standards and criteria for good work is one of the most difficult aspects of involving them in their own assessment. Again, teachers can use various ways to help students develop and cultivate these insights. Following the example of Ms. K's class in the first vignette, students and teachers can become engaged in a substantive, assessment conversation about what is a good presentation, such as a good lab investigation or a good reading summary while engaging students in the development process of assessment rubrics. Another starting point for these conversations could be a discussion about exemplary pieces of work, where students need to think about and share the characteristics of the piece of work that makes it "good."

In the first vignette, Ms. K facilitates frequent conversations with her class about what constitutes good work. Although these discussions occur at the

beginning of the project period, she regularly and deliberately cycles back to issues of expectations and quality to increase their depth of understanding as they get more involved in their projects. In discussions of an exemplary piece of work, she encourages the students to become as specific as possible. Over time, the students begin to help refine some of the criteria by which they will be evaluated. Such a process not only helps to make the criteria more useful; it increases their ownership of the standards by which judgments will be made about their work. For her third graders, Ms. R provides guidelines for planning and presenting their instruments and introduces questions for the students to address as they engage in their work.

## WHERE ARE YOU NOW?

Once they have clearly determined where they want to go, teachers and students need to find out where students currently stand in relation to the goals. Of course, the process is not quite so linear. It is not unusual for the goals to change somewhat as the students and teachers get more involved in the study.

### Variety Is Essential

Ms. K's and Ms. R's classrooms demonstrate the many ways assessment information can be obtained. In the first scenario, conferences with students allow Ms. K to ask questions, hear specifics of project activity, and probe student reasoning and thought processes. She can get a sense of how and where the individuals are making contributions to their group's work and help to ensure that they share the work at hand, including development of an understanding of the underlying processes and content addressed by the activity. The information she learns as a result of these conferences will guide decisions on time allocation, pace, resources, and learning activities that she can help provide. After observations and listening to students discuss instruments, Ms. R made the judgment that her students were ready to continue with the activity. The journals prepared by Ms. K's students and the individual reflections of Ms. R's provided the teachers with an indication of their understanding of the scientific concepts they were working with, and thereby allowed them to gain new and different insights into their respective students' work. The entries also provided the teachers with a mechanism, though not the only one, to gain some insight into the individual student's thinking, understanding, and ability to apply knowledge. In Ms. K's class, the journal writing was regular enough that the teacher's comments and questions posed in response to the entries could guide the students as they revisit previous work and move on to related activities and reflections.

Through such varied activities, the teachers in the vignettes are able to see how the students make sense of the data, the context into which they place the data, as well as the opportunity to evaluate and then assist the students on the ability to articulate their understandings and opinions in a written format or by incorporating understandings into a design. As they walk around the room,

listening, observing, and interacting with students, both teachers take advantage of the data they collect.

Any single assessment is not likely to be comprehensive enough to provide high-quality information in all the important areas so that a student or teacher can make use of the data. Ms. K, for example, would not use the student conferences to obtain all the information she needs about student comprehension and involvement. She gets different information from reading student journals. In the individual reflections, Ms. R can get additional data to complement or reinforce the information obtained by observing students as they engage in the activity or by talking with them.

## Questioning

The occasions to sit with, converse with, question, and listen to the students gave Ms. K and Ms. R the opportunities to employ powerful questioning strategies as an assessment tool. When teachers ask salient open-ended questions and allow for an appropriate window or wait time (Rowe, 1974)--they can spur student thinking and be privy to valuable information gained from the response. Questions do not need to occur solely in whole-group discussion. The strategy can occur one-on-one as the teacher circulates around the room. Effective questioning that elicits quality responses is not easy. In addition to optimal wait-time, it requires a solid understanding of the subject matter, attentive consideration of each student's remarks, as well as skillful crafting of further leading questions. In the vignette, Ms. K needed to be aware of the existence and causes of algal blooms in order to ask questions that may lead her students down productive paths in exploring them.

## Examination of Student Work

The close examination of student work also is invaluable, and teachers do it all the time. When looking at work, it is important to ask critical questions, such as "For what does this provide evidence?" "What do they mean by this response?" "What other opportunities did the child have to demonstrate knowledge or skills?" "What future experience may help to promote further development?" "What response am I expecting?" "What are the criteria for good work?" "What are the criteria for gauging competency?" These are just a few of the questions that can spur useful analysis. Continued and careful consideration of student work can enlighten both teacher and student.

## Form to Match Purpose

Like Ms. K and Ms. R in the vignettes, teachers are not concerned with just one dimension of learning. To plan teaching and to meet their students' needs, they need to recognize if a student understands a particular concept but demonstrates difficulty in applying it in a personal investigation or if a student

does not comprehend fundamental ideas underlying the concept. Specific information regarding the sources of confusions can be useful in planning activities or in initiating a conversation between students and the teacher. An array of strategies and forms of assessment to address the goals that the student and teacher have established allows students multiple opportunities to demonstrate their understandings. This is important if we hope to support all students. Darling-Hammond (1994) comments, "if assessment is to be used to open up as many opportunities as possible to as many students as possible, it must address a wide range of talents, a variety of life experiences, and multiple ways of knowing" (p. 17).

A comprehensive understanding of science requires more than knowledge of scientific information and skills. The *Standards* articulate the breadth and depth of what it means to know and be able to do in science at different grade levels. To help ensure that assessment addresses and supports a broader view of science understanding, it can be helpful to consider the different dimensions that comprise knowledge in science. Some aspects of science knowledge are highlighted in [Box 3-3](#).

With knowledge of the student's strengths, a teacher can help ensure that any particular assessment allows the student to demonstrate understanding and can assess whether information would be better gathered in a different format to allow for that opportunity to express thinking in different ways. For instance, Ms. K collects her assessment data from a variety of places, including discussions, conversations, conferences, observations, journals and written work, in addition to providing useful information, relying on a variety of sources and using a variety of formats so as not to privilege any one way of knowing. The conferences she sets up and the conversations that ensue give her opportunities to probe understandings and confusions and reach students that may not be as articulate when it comes to written work.

Thus the form that assessment takes is significant. The form and content of assessment should be consistent with the intended purpose. Underlying this guideline is the technical notion of validity. Technical features are discussed later in this chapter. Validity centers on whether the assessment is measuring or capturing what it is intended to measure or capture. If content understanding is the goal, it is necessary to design an appropriate assessment that would tap into that dimension of their understanding. If the ability to design an investigation is the goal, it is necessary to provide the opportunity for a student to demonstrate her ability to do such an activity. Validity is not, then, an inherent property of an individual assessment; rather, the interpretations drawn from the data and the subsequent actions that ensue are either valid or invalid. Choices for the form of the assessments are extensive and should be guided by the goals set for student learning. To find the direction for best use of the assessment data, a teacher or student gathers data in the course classroom activity by asking questions, such as "What does this information tell me?" and "How can I use it to further learning and improve teaching?" and "What other types of data should I be looking for to help me make sense of this information?"

From Stiggins' (2001) book, *Student-Involved Classroom Assessment*, [Figure](#)

**3-1** offers questions to consider when designing, selecting, or implementing an assessment. After first advising teachers to set clear and appropriate targets--or learning and performance goals--and convey these targets to their students, he stresses the importance of selecting appropriate methods and of taking care to avoid invalidity and bias.

### Subject-Matter Goals

Effective formative assessment must be informed by theories to ensure that it elicits the important goals of science, including a student's current understanding and procedural capability. The elements of curriculum goals and methods of instruction come together, for part of the instructor's task is to frame subgoals that are effective in guiding progress towards curriculum goals. However, this can only be done in light of the teacher's beliefs about how best to help students to learn. This introduces learning theory in addition to assessment, but in formative assessment these are very closely intertwined. Thus there has to be a conceptual analysis of the subject goals, which also is complemented by analysis of the cognitive capacities of the learners. Examples of issues that might arise are the choice between concrete but limited instances of an idea and abstract but universal presentations, the decision about whether to use daily experience or second-hand evidence, the complexity of the patterns of reasoning required in any particular approach, and research evidence about common misconceptions that hinder the progress of students in understanding particular concepts. (For additional information on these theoretical underpinnings, see NRC, 1999a.)

Here again, depth in a teacher's subject-matter knowledge is essential. When teaching the concept of force in his high school class, Jim Minstrell is aware that although students use terms like "push" and "pull" to describe "force," the understandings they have for these terms and for the concept of force differs from those shared by scientists (Minstrell, 1992). Specifically, students often believe that a push or a pull--or a force--must be due to an active, or causal, agent. With this in mind, Minstrell carefully designs his instruction, including his questions and student experiences, to help them challenge their notions as they move towards a better understanding of the scientific phenomena and explanations involved with force. After spending time discussing and drawing the forces involved as an object is dropped to the floor, he plans questions and activities to help cultivate student understandings of more passive actions of forces so they understand that the conceptual notion of force applies to both active and passive actions and objects. His class discusses the forces involved with an object resting on a table, including the reasonableness of a table exerting an upward force. They go over other situations that would help them decide what is happening in terms of force, such as discussing the forces involved as the same object sits in the student's hand, hangs from a spring, and as the object is pushed off the edge of the table. Throughout the unit, the teacher listens carefully to his students' responses and explanations. Without an understanding of both student learning and the science involved, upon hearing the proper terms from his students, he may have proceeded with his unit with the impression that the students shared a scientific understanding of



force (for a class transcript and analysis by the teacher, see Minstrell, 1992).

## Nature and Form of Feedback

The data produced from the variety of assessments illustrated in the vignettes are not only useful for the teachers but also as essential tools in helping students to realize where they stand in relation to their goals. Thus for the students, the journals with the teacher's comments added, serve as a repository for one form of feedback so they can maintain a continuing record of their work and progress. It is important to emphasize that assigning grades on a student's work does not help them to grasp what it takes on their part to understand something more accurately or deeply. Comments on a student's work that indicate specific actions to close the gap between the student's current understanding and the desired goal provide crucial help if the student takes them seriously. There is well-researched evidence that grades on student work do not help learning in the way that specific comments do. The same research shows that students generally look only at the grades and take little notice of the comments if provided (Butler, 1987). The opportunity that Ms. R's students had to design, build, and then rebuild instruments based on their trials gives them a chance to make good use of feedback to improve their piece of work.

Providing information to students is not solely a cognitive exchange. It is intertwined with issues of affect, motivation, self-esteem, self-attribution, self-concept, self-efficacy, and one's beliefs about the nature of learning. From many studies in this area (Butler, 1988; Butler & Neuman, 1995; Cameron & Pierce, 1994; Kluger & deNisi, 1996), a further generalization emerges. This is the distinction between feedback that emphasizes *learning goals* and the associated targets and feedback that focuses on *self-esteem*, often linked to the giving of grades and other reward and punishment schemes. Upon comparison of feedback in experimental studies, it is the feedback about learning goals that shows better learning gains. Feedback of the self-esteem type (trying to make the student feel better, irrespective of the quality of the work) leads less successful students to attribute their shortcomings to lack of ability. The corollary for these students is that there is little point in trying or hoping for better.

The way in which information is provided is therefore a delicate matter. Grades, and even undue praise, can reinforce expectations of failure and lead to reluctance to invest effort. Yet this culture is deeply embedded in American schools and is hard to change. This fact highlights the importance of the nature and form of the information provided to students. Thus, priority should be given to providing students with information that they can use to reach desired learning goals (Ames, 1992; Butler, 1988; Dweck, 1986).

## Timing of Assessment

In helping teachers and students establish where students stand in relation to

learning goals, assessment activities are not only useful during and at the end of a unit of teaching, they also can be valuable at the start of a piece of work. Suitably open and nontechnical questions or activities can stimulate students to express how much they already know and understand about a topic. This may be particularly important when the students come from a variety of backgrounds, with some having studied aspects of the topic before, either independently or with other teachers in different schools. Such assessment can both stimulate the thinking of the students and inform the teacher of the existing ideas and vocabularies from which the teaching has to start and on which it has to build.

### Formative Assessment in Scientific Experimentation--An Example

The following example from the Lawrence Hall of Science assessment handbook (Barber et al., 1995) demonstrates how assessment mechanisms can enrich science investigations and provide the teacher with useful information. In this illustration, students are challenged to design and conduct two experiments to determine which of three reactants--baking soda, calcium chloride, and a phenol red solution (phenol red and water)--when mixed together produces heat. The students already have completed an activity in which they mixed all three substances. The students are expected to refer to their observations and the results of that first activity. [Box 3-4](#) illustrates a data sheet used by the students for the assessment activity, which provides prompts to record their experimental design and observations. Through this investigation, the teacher would be able to assess students' abilities to do the following:

- Design a controlled experiment in which only one ingredient is omitted, so there is ONLY one difference between the preliminary reaction and the comparison reaction.
- Design experiments that will provide information to help determine which reactants are necessary to produce the heat in this reaction.
- Record their experiments, results, and conclusions using chemical notation as appropriate.
- Use experiment results and reasoning skills to draw conclusions about what causes heat. (p. 152)

These students were able to arrive at some part of what would be a correct conclusion, though the degree to which the students used logical reasoning, or supported their conclusions with data, varied widely. Many came up with a correct solution but featured a noncontrol, inadequate experimental design. In addition, the recording of results and observations was accomplished with varying degrees of clarity. Their responses, and the language they use to describe and explain observations and phenomena, suggest varying levels of understanding of the chemical and physical changes underlying the reactions. Because the assessment was designed primarily to tap scientific investigation and experimentation skills and understandings, other assessments, including perhaps follow-up questions, would be required to make inferences about their level of conceptual understanding in the chemical and physical processes

involved with these reactions.

With close examination of the student work produced in this activity, teachers were able to gain insight into abilities, skills, and understandings on which they then could provide feedback to the student. It also provided the teacher with information for additional lessons and activities on chemical and physical reactions. [Boxes 3-5](#) [[3-6](#), [3-7](#), [3-8](#)] through [3-9](#) offer samples of this type of student work along with teacher commentary.

## Creating Opportunities

Ongoing, formative assessment does not solely rely on a small-group activity structure as in the vignettes. In a whole-class discussion, teachers can create opportunities to listen carefully to student responses as they reflect on their work, an activity, or an opportunity to read aloud. In many classrooms, for example, teachers ask students to summarize the day's lesson, highlighting what sense they made of what they did. This type of format allows the teacher to hear what the students are learning from the activity and offers other students the opportunity of learning about connections that they might not have made.

In one East Palo Alto, California, classroom, the teacher asked two students at the beginning of the class to be ready to summarize their activity at the end. The class had been studying DNA and had spent the class hour constructing a DNA model with colored paper representing different nucleotide bases. In their summary, the students discussed the pairing of nucleotide bases and held up their model to show how adenine pairs with thymine and cytosine pairs with guanine. Although they could identify the parts of the model and discuss the importance of "fit," they did not connect the representative pieces to a nitrogen base, sugar, and a phosphate group. When probed, they could identify deoxyribose and the phosphate group by color, but they were not able to discuss what roles these subunits played in a DNA helix. After hearing their remarks, the teacher realized that they needed help relating the generalizations from the model to an actual strand of DNA, the phenomenon they were modeling. Regardless of the format--individual, small group, whole class, project-based, written, or discussion--teachers have the opportunity to build in meaningful assessment. These opportunities should be considered in curriculum design.

## Cultivating Student Involvement in Assessment

Student participation becomes a key component of successful assessment strategies at every step: clarifying the target and purpose of assessment, discussing the assessment methods, deliberating about standards for quality work, reflecting on the work. Sharing assessment with students does not mean that teachers transfer all responsibility to the student but rather that assessment is shaped and refined from day to day just as teaching is. For student self- and peer-assessment to be incorporated into regular practice requires cultivation

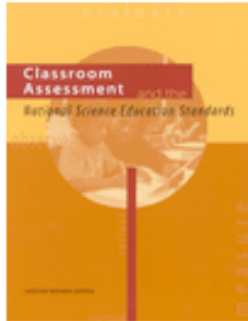
and integration into daily classroom discourse, but the results can be well worth the effort. Black and Wiliam (1998a) assert, "...self-assessment by the students is not an interesting option or luxury; it has to be seen as essential" (p. 55). The student is the one who must take action to "close" the gap between what they know and what is expected (Sadler, 1989). A teacher can facilitate this process by providing opportunities for participation and multiple points of entry, but students actually have to take the necessary action.

In the opening vignette, students in Ms. K's class are drawing on a range of data sources, including their own and classmates' projects, library research, and interviews with local experts. In preparation for presentations, the students are encouraged to make the connection of the small-scale study they do with plant fertilizer to the larger local system. Opportunities for revisions and regular discussions of what is good work help to clarify criteria as well as strengthen connections and analysis, thus improving learning. Class discussions around journal reflections provide important data for teachers about student learning and also allow students to hear connections others have made.

For this transition to occur, peer- and self-assessment must be integrated into the student's ways of thinking. Such a shift in the concept of assessment cannot simply be imposed, any more than any new concept can be understood without the student becoming an active participant in the learning. Reflection is a learned skill. Thus, the teacher faces the task of helping the student relate the desired ability to his or her current ideas about assessing one's self and others and how it can affect learning. How do students now make judgments about their own work and that of others? How accurate are these judgments? How might they be improved? Such discussions are advanced immeasurably through the examination of actual student work--initially perhaps by the examination of the anonymous work of students who are not members of the class.

Involving students in their own and peer assessment also helps teachers share the responsibility of figuring out where each student is in relation to the goals or target and also in developing a useful plan to help students bridge the gap. In addition to helping students learn how to learn, there are pedagogical payoffs when students begin to improve their ability to peer- and self-assess. Collecting and utilizing student data for every student in the classroom is made much easier with a classroom of people assisting in the same task. With a clearer vision of peer- and self-assessment and adequate time, teachers can get this help from their students and in the process help them to improve the quality of their own work.

Although there is no one way to develop peer- and self-assessment habits in students, successful methods will involve students in all aspects of the assessment process, not solely the grading after an exercise is completed. If students are expected to effectively participate in the process, they then need to be clear on the target and the criteria for good work, to assess their own efforts in the light of the criteria, and to share responsibility in taking action in the light of feedback. One method that has proved successful has been to ask students to label their work with red, yellow, or green dots. Red symbolizes the student's view that he or she lacks understanding, green that he or she has



**BUY THIS BOOK!**

confidence, and yellow that there appear to be some difficulties and the student is not sure about the quality of the response. These icons convey the same general meaning of traffic lights and are so labeled in the class. This simple method has proved to be surprisingly useful with the colored dots serving to convey at a glance, between student and teacher and between students and their peers, who has problems, where the main problems lie, which students can help one another, and so on. The traffic-light icons can play another important role, in that they help to make explicit the "big" concepts and ideas of a unit.

With a teacher's help, much useful work in student groups can start from assessment tasks: each member of a group can comment on another's homework, or one another's tests, and then discuss and defend the basis for their decisions. Such discussions inevitably highlight the criteria for quality. The teacher can help to guide the discussions, especially during the times in which students have difficulty helping one another. Peers can discuss strengths and areas of weakness after projects and presentations. Much of the success of peer- and self-assessment hinges on a classroom culture where assessment is viewed as a way to help improve work and where students accept the responsibility for learning--that of their own and of others in their community.

## HOW CAN YOU GET THERE?

Much as Ms. K and Ms. R do in the snapshots of their respective classes, captured in the vignettes, teachers continually make decisions about both the teaching and the learning going on in their classrooms. They make curricular decisions and decide on experiences they think can help further students' understandings. They decide when and how to introduce and approach a concept and determine an appropriate pace. They continually monitor levels of interest and engagement in curricular activity. They attend to the individual student, the small group, and the class as a whole. If data are collected and used to inform the teacher and student, assessment can play a significant role in all the decisions a teacher makes about what actions to take next. A focus on assessment cuts across multiple standards areas. [Box 3-10](#) shows how teaching standards seek to extend the purview of the teacher.

The teacher is able to see whether students are struggling with an activity or concept, whether they have developed fundamental understandings, whether they need to revisit a particular idea or need more practice to develop particular skills. Teachers need to understand the principles of sound assessment and apply those principles as a matter of daily routine practice.

With the knowledge gained from assessment data, a teacher can make choices. Thus, assessment serves not only as a guide to teaching methods but also to selecting and improving curriculum to better match the interests and needs of the students. According to the Assessment Standards (NRC, 1996), planning curricula is one of the primary uses of assessment data. Teachers can use assessment data to make judgments about

- the developmental appropriateness of the science content,



- student interest in the content,
- the effectiveness of activities in producing the desired learning outcome,
- the effectiveness of the selected examples, and
- the understanding and abilities students must have to benefit from the selected activities and examples. (p. 87)

Thus assessment data can be used immediately, as Ms. K does when she alters upcoming plans, and Ms. R does when she decides her students are ready to move on to the next stage of activity. The data also are useful when the teachers cover the material again the following year.

### **Assessment Should Be Consistent with Pedagogy**

For the data to be useful in guiding instructional decisions, the assessment methods should be consistent with the desired pedagogy. Thus, assessment takes into consideration process as well as outcomes and products and the instruction and activities that lead to those ends. Only if assessments in science classrooms can more closely approximate the vision of science education teaching and learning can they inform the teacher's work in trying to implement the emphasis in the *Standards* on students actively doing science.

### **Use of Assessment Data**

The extent to which any assessment data inform teaching and influence learning depends in a large part on use. Assessment-generated data do little good in the head of the teacher, in the grade book, or by failing to inform future decisions, such as selecting curricula, planning class time or having conversations with students. Teachers must use it to adapt their teaching to meet the needs of their students. In other words, just as teaching shapes assessment, assessment shapes teaching. The success of formative assessment hinges in large part on how the information is put to use.

With rich assessment data, a teacher can begin to develop possible explanations about what the difficulties might be for the student. If some pedagogical approach did not work the first time, is it likely to be more effective when repeated? Or, is some new approach required? Might other resources be provided? Setting subgoals is another strategy that is often effective. The student is encouraged to take smaller steps toward learning a particular concept or skill.

Peer instruction is another approach that can sometimes work in helping students reach a learning or performance target. If a teacher notices that one student seems to understand (for example, by displaying a green "traffic light") while another does not, the one who understands might help the one who does not. Students occasionally can assist one another because they themselves may have overcome a similar difficulty. Most all teachers use this technique from time to time during class discussion when they encourage the

entire group to help a student who clearly is having difficulty. The same principle can operate with just two students working cooperatively when one may have just figured out the desired response and can explain it to the other. Ms. R brought in sixth graders to assist her third graders while they made instruments. Even though help was provided to handle materials and supplies, the older students also could have been more vocal in the design and construction of the instruments.

## Assessment Data Management

Although teachers make assessments all the time, it is important that they develop a system for gathering data about student understanding and progress. This way, no child is overlooked and teachers can be sure that they focus on what they think are the most important learning goals and outcomes. The specific system certainly can vary, depending on a teacher's experience and preferences in gathering such information.

Relying on memory can be difficult with more than 150 students, with many activities, interactions, and observations and over the course of many months before summative evaluations call for the use of such information. One teacher might carry a clipboard while circulating around the room to record comments and observations. Each student has an index card on which to write questions or request an opportunity to speak with the teacher rather than to interrupt. Each day, the teacher observes a handful of students at work but this does not prevent the recording of information from conversations overheard in the room. This method of collecting data not only helps to organize the teaching but also serves as pertinent information when talking with parents and students. In a review of the relevant research in this area, Fuchs and Fuchs (1986) reported that student achievement gains were significantly larger (twice the effect size) when teachers used a regular and systematic method for recording and interpreting assessment data and providing feedback as compared to when they made spontaneous decisions.

In addition to making good use of the data, keeping good records of day-to-day assessments also is important for summative purposes. When meeting with parents or students, it is helpful to have notes of concrete examples and situations to help convey a point. Good records also can serve to address issues of accountability, a topic that will be discussed in the next chapter.

## THE EQUITY PRINCIPLE

The *Standards* were written with the belief that all students should be expected to strive for and to achieve high standards. According to the *Standards*, in addition to being developmentally appropriate, "assessment tasks must be set in a variety of contexts, be engaging to students with different interests and experiences, and must not assume the perspective or experience of a particular gender, racial or ethnic group" (p. 86). The

corresponding principle in classroom assessment is clear: Assessment is equitable and fair, supporting all students in their quest for high standards.

Equity issues are difficult to grapple with and arise at all levels of the education system and in all components of any program. All participants--teachers, students, administrators, curriculum developers, parents--are called upon to share the belief that all students can learn, and this premise needs to infuse all aspects of classroom life. Focusing on equity in classroom assessment is one part of the challenge.

For years, assessment has been used to sort and place students in such a way that all students do not have access to quality science programs (Darling-Hammond, 1994; Oakes, 1985, 1990). Depending on the form assessment takes and how the ensuing data are used, assessment can be a lever for high-quality science education for all rather than an obstacle. In research conducted by White and Frederiksen (1998) where students engaged in peer- and self-assessment strategies, traditionally low-attaining students demonstrated the most notable improvement.

Frequent and immediate feedback to students based on careful attention to daily activity--including student work, observations, participation in conversations and discussions--can provide teachers and students with valuable information. If this information is used in a manner that informs students about standards for improvement and how to attain them, it also can help support all students to achieve their potential.

Assessing students engaged in meaningful activities can promote equity in several other respects as well. For one, teachers can help create a setting where assessment-related activities engage students in experiences that help them synthesize information, integrate experiences, reflect on learning, and make broader connections. Through their regular journal reflections, the students in Ms. K's class reflected on their learning, making connections between their particular project and the local ecosystem. Assessments and assessment-related conversations can help make explicit to all students standards of quality work, make clearer the connections among seemingly unrelated content, concepts, and skills, and provide a scaffold for ongoing student self-assessment (Cole et al., 1999). Misunderstandings of the task or the context, misconceptions about the nature of the task, or difficulties with the language used, can be brought to light and dealt with, often by students helping one another.

Some people believe that the different roles a teacher plays with respect to assessment perpetuates inequitable treatment. In any personal relationship, few of us succeed in treating all of our acquaintances with equal consideration. We may be predisposed by their color, their gender, the way they talk, their social class, whether they respond to us in a warm or in a distant way, and much more. All teachers face such issues as they respond to their students as individuals. Formative assessment requires a close and often personal response. A student's answer to a question may seem strange or not well thought out. Sometimes such reactions may be justified, but sometimes they are prejudgments that may be unfair to the student. In particular, if a student is treated dismissively, then sees another student making a similar response

treated with respect, he may be unlikely to try again. So the first and hardest part of treating students equitably is to try to treat all students with the same respect and seriousness. In particular, the idea that everyone has a fixed IQ, that some are bright and some are not, and there is nothing one can do about it, can be very destructive of the kind of interaction necessary between teacher and student to advance learning. If a teacher really thinks in this way, it is highly probable that such an attitude will be conveyed, directly or indirectly, to the student. In the case of one pigeonholed as less "intelligent," the student might believe that this is a true judgment and therefore stop trying.

A different problem that leads to inequity in teaching is associated with problems of "disclosure," the technical label for the challenge of assuring that a student understands the context in which a question is framed and interprets the demand of the question in the way that the teacher intended. Some of these problems are associated with the language of a question or task. For example, both vocabulary and oral style differ among children so the teacher may communicate far more effectively with students from one socioeconomic or ethnic background than with those from another background. Many class questions or homework tasks are set in what are assumed to be realistic settings, often on the assumption that this will be more accessible than one set in abstract. One student's familiar setting, for example, a holiday drive in a car, may be uncommon for another family that cannot afford a car, or even a holiday. Ironically, some research has shown that questions set in "everyday" settings open up wider differences in response between students in advantaged compared with disadvantaged backgrounds than the same questions set in abstract contexts (Cooper & Dunne, 2000).

These problems of "disclosure," and the broader problems of bias in testing have been studied from many aspects in relation to summative tests, especially where these are developed and scored externally from the school. Although such external tests are not subject to the risks of bias at a personal, one-on-one level, this advantage may be offset because a teacher might see that a student does not understand a question and can rephrase to overcome the obstacle, the external grader or machine cannot.

Some people caution against complications associated with the multiple roles that teachers play in assessment, including that of both judge and jury. They see this subjectivity as a threat to the validity of the assessment. They point to a study that examined the effects of expectations on human judgment (Rosenthal & Jacobsen, 1968). Teachers were provided contrived information that a handful of students showed exceptional promise, when in actuality they were no different from the others. When questioned several months later about those students' progress, the teacher reported that they excelled and progressed more than their classmates. One of the basic claims made by the researchers in this study was that the teacher fulfilled the "exceptional-promise" expectation. In efforts to try to overcome or at least abate inherent bias that results in inequitable treatment, teachers, and all those working with students, need to be examined and keep a check on the bias that enters into their own questioning, thinking, and responses.

## VALIDITY AND RELIABILITY

To some, issues of validity and reliability are at the heart of assessment discussions. Although these considerations come into play most often in connection with large-scale assessment activities, technical issues are important to consider for all assessments including those that occur each day in the classroom (American Educational Research Association, American Psychological Association, & National Council on Measurement and Education, 1999). Though principles stay the same, operationally they mean and look different for formative and summative purposes of assessment.

Issues of validity center on whether an assessment is measuring or capturing what is intended for measure or capture. Validity has many dimensions, three of which include content validity, construct validity, and instructional validity. Content validity concerns the degree to which an assessment measures the intended content area. Construct validity refers to the degree to which an assessment measures the intended construct or ability. For example, the *Standards* outline the abilities and understandings necessary to do scientific inquiry. For an assessment to make valid claims about a student's ability to conduct inquiry, the assessment would need to assess the range or abilities and understandings comprised in the construct of inquiry. Finally, an assessment has instructional validity if the content matches what was actually taught. Questions concerning these different forms of validity need to be addressed independently, although they are often related. Messick (1989) offers another perspective on validity. His definition begins with an examination of the *uses* of an assessment and from there derives the technical requirements. Validity, as he defines it, is "an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and *appropriateness of inferences and actions* based on test scores or other modes of assessment" [italics added] (p. 13). Thus, validity in his view is a property of consequences and use rather than of the actual assessment. Messick's (1994) use of validity stresses the importance of weighing social consequences: "Test validity and social values are intertwined and that evaluation of intended and unintended consequences of any testing is integral to the validations of test, interpretation and use" (p. 19). Validity, he argued, needs evidentiary grounding, including evidence of what happens as a result. Moss (1996) urges that actions taken based on interpretation of assessment data and that consequences of those actions be considered as evidence to warrant validity.

Attention to issues of validity is important in the type of ongoing classroom assessment discussed thus far in this chapter. It is important to keep in mind the guideline that says that assessments should match purpose. When gathering data, teachers and students need to consider if the information accurately represents what they wish to summarize, corresponds with subject matter taught, and reflects any unintended social consequences that result from the assessment. Invalid formative assessment can lead to the wrong corrective action, or to neglect action where it is needed. Issues relating to validity are discussed further in [Chapter 4](#).

Reliability refers to generalizability across tasks. Usually, it is a necessary but



not complete requirement for validity. Moss (1996) makes a case that reliability is not a necessity for classroom assessment. She argues for the value of classroom teachers' special contextualized knowledge and the integrative interpretations they can make. Research literature acknowledges that over time, in the context of numerous performances, concerns of replicability and generalizability become less of an issue (Linn & Burton, 1994; Messick, 1994). Messick states that dropping reliability as a prerequisite for validity may be "feasible in assessment for instructional improvement occurring frequently throughout the course of teaching or in appraisals of extensive portfolios" (p. 15).

For formative assessments, constraints on reliability are handled differently though still important to consider (William & Black, 1996). If assessment takes place all the time, a teacher can elicit information that suggests that a previous assessment and judgment was not representative of performance. Teachers are in the position of being able to sample student performance repeatedly over time, thus permitting assessment-based judgments to be adjusted and evolve over a long period of time, leading to confident conclusions.

Teachers, however, must remain open to continually challenging and revising their previously held judgments about student performance. Research suggests that teachers often look for evidence that affirms their own performance (Airasian, 1991) and do not easily modify judgments on individual student achievement (Goldman, 1996; Rosenbaum, 1980).

Although teachers do have a "special-observer" perspective from which they have access to information not generated by way of a test, consideration of technical criteria should remind teachers that careful documentation and systematic observation of all students is necessary to achieve an equitable classroom environment. Assessment data should be "triangulated," or drawn from multiple sources, to reduce the possible bias that may be introduced by any one particular method of obtaining and interpreting evidence.

### Thinking in Terms of the Classroom

Thus far, this chapter has provided a menu of strategies and principles for teachers to consider when designing and implementing a classroom assessment system organized around the goals of improved student work. As noted previously, no one system or collection of strategies will serve all teachers. When choosing among the many available assessment approaches, the following general selection guidelines may be of use. For one, assessments should be aligned with curricular goals, and should be consistent with pedagogy. Because a single piece of work or performance will not capture the complete story of student understanding, assessments should draw from a variety of sources. On a related note, students should be provided with multiple opportunities to demonstrate understanding, performance, or current thinking. Assessments can be most powerful when students are involved in the process, not solely as responders or reactors. Also when designing and selecting assessment, a teacher should consider his or her personal style. Lastly, assessments should be feasible. With large class sizes and competing

priorities, some teachers may find it impractical to employ certain practices.

Although any classroom activity can be modified to also serve as an assessment, the data must be fed back into teaching and learning for the assessment to be effective. To the extent that a teacher's decisions and judgments are informed by the information they glean from their students--for example, through observations, class discussions, conversations, written comments, reflections, journals, tests, quizzes, and presentations--teachers can base decisions on understandings of their students and significantly support their learning.

Unfortunately, there are often competing needs and demands on teachers. Teachers have little choice but to juggle the different purposes of assessment in effort to create some coherent system that can best satisfy the different, and often competing, assessment aims. Because they are stretched thin with resources and time, teachers need support in helping them realize the potential of this type of assessment. We turn to this challenge in [Chapters 5](#) and [6](#).

## KEY POINTS

- To be effective as assessment that improves teaching and learning, the information generated from the activity must be used in such a way as to inform the teacher and/or her students in helping decide what to do next.
- It is important for teachers to have clear performance criteria in mind before they assess student work and responses. These should be conveyed to students.
- Form and content of assessment should be consistent with the intended purpose.
- Student participation becomes a key component of successful assessment strategies at every step. If students are expected to effectively participate in the process, then they need to be clear on the target and the criteria for good work, to assess their own efforts in light of the criteria, and to share responsibility in taking action in light of feedback.
- Assessments should be equitable and fair, supporting all students in their quest for high standards. Thus, technical issues are important to consider for all assessments, including those that occur each day in the classroom.

